PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/12, 15/11, 15/62, 5/10, C07K 14/715, 16/28, C12Q 1/68, 1/70, G01N 33/68

A1

(11) International Publication Number: **WO 98/54317**

(43) International Publication Date: 3 December 1998 (03.12.98)

(21) International Application Number: PCT/EP98/03437

(22) International Filing Date: 29 May 1998 (29.05.98)

(30) Priority Data:
60/048,057 30 May 1997 (30.05.97) US

(71) Applicant: FONDATION MONDIALE RECHERCHE ET PREVENTION SIDA [FR/FR]; 1, rue Miollis, F-75732 Paris Cedex 15 (FR).

(72) Inventors: BERETTA, Alberto; 3, avenue R. Schumann, F-75007 Paris (FR). QUILLENT, Caroline; 33, rue Clément Perrot, F-94400 Vitry sur Seine (FR). ARENZANA SIESDEDOS, Fernando; 18, rue de Rushemoor, F-92190 Meudon (FR). BRAUN, Joséphine; 5, place du Marché Sainte Catherine, F-75004 Paris (FR).

(74) Agent: CABINET ORES; 6, avenue de Messine, F-75008 Paris (FR).

(81) Designated States: CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: HUMAN IMMUNODEFICIENCY VIRUS CO-RECEPTOR VARIANTS ASSOCIATED WITH RESISTANCE TO VIRUS INFECTION

(57) Abstract

Variants of the CCR5 human immunodeficiency virus type 1 (HIV-1) co-receptor which results in a resistance of CCR5-expressing cells to HIV-1 infection. The detection of such mutations may be used to identify individuals at lower risk for infection relative to the general population who, if infected, may exhibit slower progression to AIDS. Further, the present invention provides for oligonucleotide primers and probes and diagnostic methods for detecting the presence of such variants and for methods of inhibiting HIV-1 infection of a cell expressing the CCR5 receptor, comprising introducing, into the cell, a nucleic acid encoding a CCR5 variant.

A03

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

HUMAN IMMUNODEFICIENCY VIRUS CO-RECEPTOR VARIANTS ASSOCIATED WITH RESISTANCE TO VIRUS INFECTION

The present invention relates to variants of the CCR5 human immunodeficiency virus type 1 (HIV-1) co-receptor which results in a resistance of CCR5-expressing cells to HIV-1 infection. The detection of such mutations may be used to identify individuals at lower risk for infection relative to the general population who, if infected, may exhibit slower progression to AIDS. Further, the present invention provides for oligonucleotide primers and probes and diagnostic methods for detecting the presence of such variants and for methods of inhibiting HIV-1 infection of a cell expressing the CCR5 receptor, comprising introducing, into the cell, a nucleic acid encoding a CCR5 variant.

Although human immunodeficiency virus type 1 (HIV-1) uses the T cell surface molecule CD4 as a primary receptor, successful viral entry into and infection of a cell has been found to require the presence of a second molecule, or co-receptor (Clapham et al., 1997, Nature, 388, 230-231). Seven co-receptor molecules have been identified, each of which members of, or related to the family of chemokine receptors, which are G-protein coupled receptors having seven transmembrane domains.

Chemokines are protein having molecular weights from about 7-16 kDa which, acting as ligands at chemokine receptors, induce a rapid calcium influx and mediate a number of effects on the immune system (Murphy, Cytokine Growth Factor Rev., 1996, **7**, 47-64). Examples of chemokines include macrophage inflammatory protein (MIP)-1a and MIP-1b, a protein which is regulated on activation normally T cell expressed and secreted (RANTES), monocyte chemoattractant protein (MCP)-1, MCP-2, MCP-3, MCP-4,

eotaxin and stromal-derived factor (SDF)-1 (Clapham et al., 1997, cited). Chemokines are proteins that are classified into two groups based on the presence of a non-cysteine amino acid (X) between the first two CC of four cysteine residues appearing in their amino acid sequence, giving rise to the CXC (a) family and the CC (b) family. Receptors which specifically recognize CXC or CC chemokines are referred to, accordingly, as CXCR or CCR (Dynamics of HIV infection, Science and Medicine, 1998, 36-45).

Two species of chemokine receptors which appear to be particularly relevant to HIV infection are CCR5 and CXCR4, for which the natural ligands are MIP-1a, MIP-1b and RANTES (CCR5) and SDF-1 (CXCR4). To date, most of HIV-1 clinical isolates studied appear to use CCR5 or CXCR4, or both, as co-receptors with CD4 for entry into cells, and the presence of chemokine ligand inhibits infection via the corresponding receptor.

The cellular distributions of CCR5 and CXCR4 are associated with the role of these molecules in the course of HIV-1 infection. CCR5, which is mainly expressed on macrophages and memory T cells, serves as a co-receptor for infection by macrophage-tropic (M-tropic) strains of HIV-1, which are found throughout the course of infection, are preferentially involved in sexual transmission of HIV-1 and are expressed by non-syncitium-inducing laboratory isolates which do not cause cell/cell fusion in T cell lines (Cocchi et al., Science, 1995, 270, 1811-1815; Alkhatib et al., Science, 1996, 272, 1955-1958; Choe et al., Cell, 1996,, 85, 1135-1148; Deng et al, Nature, 1996, 381, 661-666; Doranz et al, Cell, 1996, 85, 1149-1158; Dragic et al., Nature, 1996, 381, 667-673). CXCR4, however, which is expressed on a broader spectrum of cells, including naive T cells, serves as the

co-receptor in late stages of infection for syncitium-inducing T cell tropic (T-tropic) strains of HIV-1 (Belul et al., Nature, 1996, 382, 829-833; Oberlin et al., Nature, 1996, 382, 833-835; Feng et al., Science, 1996, 272, 872-877). Accordingly, the co-receptor which is more relevant to the initiation of HIV-1 infection appears to be CCR5.

Indeed, an association has been drawn between rare individuals who remain persistently uninfected despite multiple sexual exposures to HIV and the presence of a 32 base pair deletion in the CCR5 gene (CCR5Δ32: Samson et al., Nature, 1996, 382, 722; Liu et al., Cell, 1996, 367) which results in a protein having a truncation of C-terminal residues 188-352 (including the fifth, sixth and seventh transmembrane domains). Individuals heterozygous for this deletion, are however, susceptible, to infection (Dean et al., Science, 1996, 273, 1856), although progression to AIDS may be slowed (Dean et al., 1996, cited; Samson et al., 1996, cited, Huang et al., Nature Med., 1996, 2, 1240-1243, Michael et al., Nature Med., 1997, 3, 338-340). It has been proposed (Benkirane et al., J. Biol. Chem., 1997, 272, 30603-30606) that coexpression of the CCR5Δ32 gene with wild type CCR5 gene results in trans-inhibition of the ability of CCR5 to act as an HIV co-receptor, in which the CCR5Δ32 protein interferes with dimerization of CCR5 at the cell surface. It has not, however, been confirmed that the dimerization of CCR5 occurs or is necessary for viral entry.

The present invention relates to truncated variants of CCR5 which lack the portion of the molecule beginning with the third transmembrane domain. It is based, at least in part, on the discovery of a specific variant form of CCR5, termed CCR5m303, in which the gene

is mutated and creates a stop codon which arrests translation before the third transmembrane domain.

Unexpectedly, the presence of both CCR5m303 and CCR5Δ32 variant alleles in individuals was observed to confer resistance to infection by M-tropic strains of HIV-1.

The present invention is also based on the fact that the CCR5m303 variant is unexpectedly more effective than CCR5Δ32 in trans-inhibiting the ability of wild type CCR5 to act as a co-receptor for HIV. Therefore, individuals having a genotype which includes a wild type CCR5 allele and a CCR5m303 variant allele may be protected against infection by M-tropic strains of HIV.

Accordingly, in a first series of embodiments, the present invention provides for compositions comprising a nucleic acid encoding the CCR5m303 variant and portions thereof which contain or which may be used to detect the m303 mutation.

In a second series of embodiments, the present invention provides for methods of identifying the presence of the CCR5m303 variant in an individual, wherein such methods may also include the identification of a second species of the CCR5 variant. The presence of the CCR5m303 variant, in conjunction with wild type CCR5 or a second species of CCR5 variant, bears a positive correlation with resistance to infection with M-tropic strains of HIV-1 and may be indicative of slower progression of disease in heterozygous individuals.

In a third set of embodiments, the present invention provides for compositions comprising a nucleic acid encoding a CCR5 variant which comprises the first two transmembrane domains found in wild type CCR5 but lacks the remainder of the C-terminal end of the mole-

cule, and for the corresponding CCR5 variant proteins.

In a fourth set of embodiments, the present invention provides for methods of inhibiting CCR5-mediated HIV infection of a cell comprising decreasing the number of functional CCR5 molecules present at the surface of the cell. Such methods include, but are not limited to, introducing, into the cell, a nucleic acid encoding a CCR5 variant which comprises the first two transmembrane domains found in wild type CCR5 but lacks the remainder of the C-terminal end of the molecule.

Figure 1A-D. Infection of peripheral blood mononuclear cells (PBMC) from two unexposed uninfected (UU) and two exposed uninfected (ExU) individuals with CCR5-dependent (YU2) and CXCR4-dependent (NL4-3) HIV molecular clones after inoculation with 20 ng p24/10^6 cells, with or without addition of RANTES (R). Results are expressed as the amount of p24 antigen in culture supernatants, and are representative of four independent experiments. CCR5 genotype for the Δ32 deletion is indicated in parenthesis (wt/wt) = homozygous wild type; (wt/Δ)= heterozygous Δ32; (Δ/Δ)= homozygous Δ32; and wt?/Δ)= heterozygous Δ32 with an apparent wild type allele.

Figure 2. Functional analysis of both CCR5 alleles cloned from ExU2 using an *Env*-mediated cell fusion assay.

Figure 3A-D. Genetic analysis of CCR5 from two UU individuals, two ExU individuals, and two siblings of ExU2. (A) DNA nucleotide and amino acid sequences of wild type (wt) and m303 alleles in a region spanning the 303 mutation (SEQ ID NO:11 and SEQ ID NO:12). (B) PCR amplification of genomic DNA using specific primers for the Δ32 deletion (C) PCR amplification of the entire CCR5

gene from genomic (lanes 1-6) and plasmid (lanes 7-9) DNA, followed by (D) HincII digestion after amplification. A 1kb DNA ladder was used as a marker (M). Genomic DNA following Δ32 PCR analysis is designated as follows : ExU2 : Exposed Uninfected (m303/Δ32) ; 2.S : sister of ExU2 (wt?/Δ) 2.F : father of ExU2 (wt?/Δ) ; UU1 : Unexposed Uninfected (wt/wt) ; ExU1 : Exposed Uninfected (Δ32/Δ32) ; UU2 : Unexposed Uninfected (wt/Δ) ; Plasmid DNA corresponds to wild type (wt), Δ32, and m303 cloned alleles. wt? : apparent wild type allele.

Figure 4A-B. Infection of PBMC from family members of ExU2 with CCR5-dependent (YU2) and CXCR4-dependent (NL4-3) viruses. Results are expressed as the amount of p24 antigen in cell-free culture supernatants. Day 12 of infection of PBMC from ExU2, 2.S, 2.F and UU2 with 0.5 ng p24/10^6 cells of (A) YU2 or (B) NL4-3 HIV molecular clones.

Figure 5A-B. Cotransfection experiments

The present invention relates to variants of CCR5 which comprise the first two transmembrane domains but lack the remainder of the C-terminal portion of the molecule. Such variants lack transmembrane domains 3-7 as found in wild type CCR5.

These transmembrane domains are positioned hereafter in reference to sequence SEQ ID NO:12, which corresponds to the sequence of the CCR5 human gene (wild type), in which XXX means TGT and YYY means cysteine:

- transmembrane domain 1 (TM1) corresponds to the nucleic fragment 333-407 of SEQ ID NO:12 and to the protein fragment 32-56 of SEQ ID NO:13;

- transmembrane domain 2 (TM2) corresponds to the nucleic fragment 438-500 of SEQ ID NO:12 and to the protein fragment 67-87 of SEQ ID NO:13;

- transmembrane domain 3 (TM3) corresponds to the nucleic fragment 546-611 of SEQ ID NO:12 and to the protein fragment 103-124 of SEQ ID NO:13;

- transmembrane domain 4 (TM4) corresponds to the nucleic fragment 663-740 of SEQ ID NO:12 and to the protein fragment 142-167 of SEQ ID NO:13;

- transmembrane domain 5 (TM5) corresponds to the nucleic fragment 837-908 of SEQ ID NO:12 and to the protein fragment 200-223 of SEQ ID NO:13;

- transmembrane domain 6 (TM6) corresponds to the nucleic fragment 945-1019 of SEQ ID NO:12 and to the protein fragment 236-260 of SEQ ID NO:13;

- transmembrane domain 7 (TM7) corresponds to the nucleic fragment 1062-1142 of SEQ ID NO:12 and to the protein fragment 275-301 of SEQ ID NO:13.

With reference to the amino acid sequence of wild type human CCR5 (SEQ ID NO:12), the present invention relates to CCR5 variants which comprise a portion having amino acid residues 1-101 of SEQ ID NO:13; said sequence is represented at SEQ ID NO:14 and SEQ ID NO:15 and spans the first two transmembrane domains of wild type CCR5 ; it therefore lacks the transmembrane domains 3-7 as specified hereabove.

Accordingly, the present invention provides for a purified CCR5 variant protein which comprises the first two transmembrane domains of wild type CCR5 but not transmembrane domains 3, 4, 5, 6, and 7 (SEQ ID NO:14 and SEQ ID NO:15). The CCR5 variant may further comprise additional heterologous amino acids as a fusion protein.

The present invention also provides for nucleic acid molecules encoding such CCR5 variant proteins. Such nucleic acids may themselves be truncated or may comprise a termination codon at positions 540, with reference to SEQ ID NO:12.

In a specific non-limiting embodiment, the present invention provides for a purified protein which is the CCR5 variant CCR5m303, having an amino acid sequence as set forth in SEQ ID NO:15 and for fusion proteins comprising the CCR5m303 variant joined to a heterologous protein sequence. The present invention also provides for purified and isolated nucleic acid molecules encoding such CCR5 variant proteins, including a nucleic acid having a sequence as set forth in SEQ ID NO:14.

The proteins of the invention may be prepared by synthetic techniques (Merrifield method, for instance), by cleavage of naturally derived CCR5 or, preferably, by recombinant techniques.

The present invention also includes nucleotide fragments, in particular oligonucleotides, derived from nucleotide sequences such as defined above, and in particular the following oligonucleotides, which are useful as primers :

- SEQ ID NO:1 (**Δ32 sens**) : GTCTTCATTA CACCTGCAGC TC
- SEQ ID NO:2 : (**Δ32 inv**) : GTGAAGATAA GCCTCACAGC C
- SEQ ID NO:3: **(SE)** : CCCAAGCTTA TGGATTATCA A
- SEQ ID NO:4 : **(ASE)** : GCTCTAGATC ACAAGCCCAC AGA
- SEQ ID NO:5 **(S1)**: GGGCAACTAA ATACAT
- SEQ ID NO:6: **(S2)**: GCACAACTCT GACTGG
- SEQ ID NO:7: **(mD1)**: TCC TTC TTA CTG TCC CCT TCT GG
- SEQ ID NO:8: **(mD2)**: CCT GTG CCT CTT CTT CTC ATT TC

The position of these different primers on the gene sequence are the following:

SEQ ID NO:	Primer	Position on the gene sequence
1	D32sens	757
2	D32inv	955
3	SE**	240
	cccaagctt**atggattatcaagtg** bold sequence=CCR5 sequence	
4	ASE*	1298
	gctctaga**tcacaagcccacaga** bold sequence=CCR5 sequence	
5	S1	173
6	S2	1338
7	md1	475
8	md2	935

**SE primer includes a restriction site HindIII

* ASE primer includes a restriction site XbAI.

The present invention also includes nucleotide fragments, useful as primers, which comprises between 15 and 100 nucleotides of a sequence upstream or downstream nucleotide 540 (with reference to the wild type human CCR5 gene).

The invention also relates to nucleotide fragments which are complementary to those above.

The present invention also includes fragments which may be useful as probes ; among such fragments, one may cite SEQ ID NO:10 and SEQ ID NO:11.

The nucleic acids of the invention may be incorporate into suitable vectors for cloning and/or for expression, and as such, may be operatively linked to appropriate promoter sequences, ribosome binding sequences, signal sequences, transcription termination sequences polyadenylation sequences, splice donor/acceptor sequences etc... Examples of such vectors include but are not limited to, herpes simplex viral based vectors such as pHSV1 (Geller et al., Proc. Natl.

Acad. Sci., 1990, 87, 8950-8954); retroviral vectors such as MFG (Jaffee et al., Cancer Res. 53, 2221-2226, 1993,), and in particular Moloney retroviral vectors such as LN, lNSX, LNCX, LXSN (Miller and Rosman, Biotechniques 7, 980-989, 1989); vaccinia viral vectors such as MVA (Sutter and Moss, Proc. Natl. Acad. Sci. U.S.A., 89, 10847-10851, 1992); adenovirus vectors such as pJM17 (Ali et al., Gene Therapy 1, 367-384, 1994; Berker, Biotechniques 6, 616-624, 1988; Wand and Finer, Nature Medicine, 2, 714-716, 1996); adeno-associated virus vectors such as AAV/neo (Mura-Cacho et al., J. Immunother. 11, 231-237, 1992); lentivirus vectors (Zufferey et al., Nature Biotechnology, 15, 871-875, 1997); baculovirus expression vectors such as p2Bac, and plasmid vectors such as pcDNA3 and pcDNA1 (Invitrogen), pET 11a, pET3a, pET11d, pET3d, pET22d, pET12a and pET28a (Novagen) ; plasmid AH5 (which contains the SV40 origin and the adenovirus major late promoter), pRC/CMV (Invitrogen), pCMUII (Paabo et al., EMBO J. 5, 1921-1927, 1986), pZipNeo SV (Cepko et al., Cell 37, 1053-1062, 1984), pSRα (DNAX, Palo Alto, CA) and pBK-CMV, pSPTg-T2FpAXK and pSPTg.2FXK (Schaleger et al., Proc. Natl. Acad. Sci. U.S.A., 94, 3058-2063, 1997).

Suitable expression systems include mammalian cells, insect cells, yeast, bacteria, and plants.

The present invention also provides for a cell into which any of the foregoing nucleic acids has been introduced. The cell may be a vertebrate cell such as a mammalian cell (including, but not limited to, a human cell), a bacteria cell, a yeast cell, a plant cell, or an insect cell. The nucleic acid may be introduced by transfection, injection, electroporation, transformation, cell fusion, or any other standard technique.

The CCR5 variants of the invention either

prevent or inhibit infection of a cell by a HIV virus which requires CCR5 as a co-receptor. The ability of the CCR5 variants to block or inhibit infection may be confirmed in an *Env*-mediated cell fusion assay such as the CD4+ HeLa LTR/*lacZ* assay described hereafter in the examples.

The present invention provides for a method of detecting the presence of a CCR5 variant which comprises the first two transmembrane domains (1 and 2) of wild type CCR5 but which lacks transmembrane domains 3, 4, 5, 6 and 7 in an individual, wherein the presence of such a variant bears a positive correlation with either a resistance to infection by M-tropic strains of HIV-1, a lower probability of infection or a slower progression of disease if such infection has occurred.

In particular, a determination that both CCR5 alleles of an individual encode a CCR5 variant which either (i) comprises the first two transmembrane domains of wild type CCR5 but which lacks transmembrane domains 3, 4, 5, 6 and 7 or (ii) is functionally defective (e.g., cannot be expressed at the cell surface) has a positive correlation with a resistance of the individual to infection by M-tropic strains of HIV-1.

One non-limiting example of a functionally defective CCR5 variant is the CCR5Δ32 variant.

Further, a determination that one CCR5 allele of an individual is a wild type CCR5 allele and the other allele is a CCR5 variant which comprises the first two transmembrane domains of wild type CCR5 but which lacks transmembrane domains 3, 4, 5, 6 and 7 may have a positive correlation with a decreased risk of infection, relative to homozygously wild type individuals, and with slower progression of disease in the individual once infection with an M-tropic strain of HIV-1 has occurred.

The phrase "slower progression of disease", as used herein, refers to a longer time interval between infection and progression to a diagnosis of acquired immunodeficiency syndrome (AIDS).

In preferred embodiments, the methods of the invention are directed towards detecting the presence of the CCR5m303 variant in an individual.

Non-limiting examples of genotypes having a positive correlation with resistance to infection to macrophage-tropic strains of HIV-1 would include individuals homozygous for CCR5m303, and individuals having one CCR5m303 allele and one CCR5Δ32 allele.

Non-limiting example of genotypes having a positive correlation with a lower probability of infection with macrophage-tropic strains of HIV-1 or a slower progression of disease once infection had occurred would include individuals having one wild type CCR5 allele and one CCR5m303 allele.

The presence of a CCR5 variant, as set forth above, or of wild type CCR5, may be accomplished by detecting the presence of a nucleic acid or a protein which is characteristic of said variant or wild type CCR5.

The instant invention also relates to a method of detecting and identifying the presence of at least a CCR5 variant in a biological sample, comprising :

- bringing into contact the biological sample with a pair of primers selected from the group consisting of:

SEQ ID NO:1 and SEQ ID NO:2
SEQ ID NO:3 and SEQ ID NO:4
SEQ ID NO:5 and SEQ ID NO:6
SEQ ID NO:7 and SEQ ID NO:8

- amplifying the CCR5 variant, eventually present in said biological sample,

- demonstrating the amplification of the DNA fragment corresponding to the fragment flanked by the primers, for example by gel electrophoresis,

- optionally, verifying the sequence of the amplified fragment, for example by specific probe hybridization, sequencing or restriction site analysis.

Nucleic acid and amino acid sequences of wild type CCR5 are set forth herein as SEQ ID NO:12, wherein XXX = TGT and YYY = Cys or SEQ ID NO:13. Nucleic acid and amino acid sequences of the CCR5m303 variant are set forth herein as SEQ ID NO:14 or SEQ ID NO:15. The CCR5Δ32 variant is characterized by the nucleic acid set forth in SEQ ID NO:16, having a deletion of nucleic acid residues [position 793-824] and an amino acid sequence consisting essentially of residues 1-187 of SEQ ID NO:13.

Analysis may be performed using a suitable sample collected from the individual, including, but not limited to, a blood sample, and in particular of PBMC.

The presence of the CCR5m303 variant or CCR5Δ32 variant in an individual may be detected by obtaining and sequencing the CCR5 alleles of the individual, or portions thereof spanning the 303 position or the Δ32 position, using standard techniques which preferably employ amplification technology, with the primers and the probes hereabove defined.

Using such technology, a biological sample collected from a subject to be tested is contacted with a pair of oligonucleotide CCR5-directed primers, under conditions which allow for the hybridization of the primers to nucleic acid template in the sample, the primers are extended under suitable conditions,

dissociated from the template, and then re-annealed, extended, and dissociated such as to amplify the number of copies of CCR5 nucleic acid demarcated by the primers, and then the product of amplification may be characterized by electrophoresis, restriction endonuclease cleavage patterns, and/or nucleic acid sequencing using standard techniques.

In particular preferred non-limiting embodiments of the invention, the presence of the CCR5m303 allele in an individual may be detected by determining that a CCR5 allele in the individual lacks of HincII restriction enzyme cleavage site at the 303 position.

Nucleic acid encoding a CCR5 allele, prepared from a subject (e.g., from PBMC), may be tested for the presence or absence of this cleavage site, either directly, or, preferably, by amplification techniques. For example, the entire CCR5 gene sequence, or a portion thereof spanning the 303 site, may be amplified using the primers specified above.

In one non-limiting specific example, for obtaining amplified DNA encoding the entire CCR5 gene, genomic DNA may be amplified using SEQ ID NO:3 and SEQ ID NO:4 or SEQ ID NO:5 and SEQ ID NO:6.

The restriction endonuclease HincII may then be added to amplified DNA prepared from the subject being tested (and preferably, in parallel, to a corresponding DNA sample representing the wild type CCR5 gene), under conditions recommended by the enzyme manufacturer for cleavage. The reaction product may then be separated to reveal restriction fragments, for example by electrophoresis in a 1 percent agarose gel. The absence of a HincII cleavage site at the 303 position is indicative of the presence of the CCR5m303 mutation. Where the entire CCR5 sequence is analyzed in this matter, the CCR5m303 muta-

tion is associated with a single restriction fragment approximately 1.1 kb in size, compared with two fragments of approximately 0.7 and 0.4 kb in wild type and CCR5Δ32 alleles (the HincII site between the fragments evidently being eliminated by the mutation).

Identification of the presence of the CCR5Δ32 variant, in a specific non-limiting embodiment, may be performed using the primers SEQ ID NO:1 and SEQ ID NO:2 or SEQ ID NO:7 and SEQ ID NO:8.

The present invention also includes a method of rapid detection of the presence of two simultaneous variants of the CCR5 gene, wherein the first variant is a CCR5m303 variant and the second variant is a CCR5Δ32 variant, wherein it comprises :

a step in which the biological sample is brought into contact with the pair of primers including SEQ ID NO:7 and SEQ ID NO:8 and

a step in which the product resulting from the interaction between the nucleotide sequence of the CCR5m303 allele and the CCR5Δ32 allele may be detected by any suitable means (probe hybridization, sequencing or restriction site analysis).

Identification of the foregoing variants may also be performed at the protein level, for example, by subjecting a protein sample collected from an individual and subjecting such protein to Western blot analysis, wherein an antibody directed against CCR5 is used to identify CCR5 proteins expressed in the individual. The appearance of a CCR5 protein having a molecular weight which is lower than wild type [40.6 kDa], and preferably of about [13-21.5 kDa], is indicative of the presence of a CCR5 variant having the characteristics set forth above.

As such, the present invention provides for molecules, compositions and kits which may be used in the foregoing analysis. Such molecules include but are not limited to oligonucleotide molecules which may be used to detect nucleotide defects in a CCR5 gene which give rise to the variants describe above. Such oligonucleotides preferably have a length between 8 and 100 and more preferably between 18 and 50 bases in length, and may optionally be detectably labeled, for example, with a radioactive or a non-radioactive compound. Specific non-limiting embodiments are oligonucleotides comprising the sequences SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:10 and SEQ ID NO:11, wherein SEQ ID NO:1 and SEQ ID NO:2 are particularly useful in detecting the presence of CCR5Δ32; SEQ ID NO:5 and SEQ ID NO:6 are useful in amplifying the entire CCR5 gene and SEQ ID NO:7 and SEQ ID NO:8 are useful in detecting the presence of both CCR5Δ32 and CCR5m303.

Suitable kits for performing methods of the invention may comprise (i) at least a CCR5 oligonucleotide primer and (ii) a reagent for performing an amplification reaction therewith. The present invention further provides for oligonucleotide primers (SEQ ID NO:7 and SEQ ID NO:8) designed to amplify a fragment of the CCR5 gene spanning the 303 position but not comprising the entire gene; the region spanned may preferably be at least 0.5 kb in length.

Because it has been observed that the CCR5m303 variant has a trans-inhibitory effect on the ability of wild type CCR5 to act as a suitable co-receptor for M-tropic strains of HIV-1, the present invention provides for methods of inhibiting infection of a CCR5-expressing cell comprising introducing, into the cell, a CCR5

variant which comprises the first two transmembrane domains (1 and 2) of wild type CCR5 but which lacks transmembrane domain 3, 4, 5, 6 and 7. In preferred embodiments, the CCR5 variant is CCR5m303.

The CCR5 variant may be introduced into the cell by introducing, into the cell, a nucleic acid encoding the CCR5 variant, either by infection with a virus containing the CCR5-encoding nucleic acid (such as a retrovirus, adenovirus, adeno-associated virus, etc.) or by injection of nucleic acid into the cell or surrounding tissue. Such introduction may be accomplished when the cell is part of a subject to be treated (an *in vivo* approach) or when the cell has been removed from the subject (an *ex vivo* approach).

As a specific, non-limiting example, nucleic acid encoding the CCR5m303 variant may be introduced selectively into cells which are CCR5+ and CD4+. In further specific non-limiting embodiments, expression of the nucleic acid encoding the CCR5 variant may be placed under the control of an inducible promoter or such that a trans-acting sequence is required (for example by an HIV-encoded factor, such as *tat*, so that expression in the cell is not constitutive.

EXAMPLE

MATERIALS AND METHODS

Determination of susceptibility of PBMC from UU or ExU individuals to infection, and of effects of RANTES on infection.

Cultures of peripheral blood mononuclear cells (PBMC) collected from two unexposed uninfected (UU) and two exposed uninfected (ExU) individuals were inoculated with either CCR5-dependent (YU2) or CXCR4-dependent (NL4-3) HIV molecular clones, using a virus inoculum of 20 ng

p24/10^6 cells, with or without addition of RANTES (R), as follows. Ficoll purified human PBMC were activated by 1 µg/ml PHA-P (Glaxo Wellcome, Paris, France) for 72 h in RPMI 1640 medium (Gibco/BRL Life Technologies, Cergy, France) containing 10% fetal calf serum (FCS) and infected with HIV-1 containing an amount of p24 antigen determined by ELISA (Dupont de Nemours, Wilmington, DL) for 2 h. Cells were washed three times with PBS and cultured at 10^6 cells/ml in RPMI 1640 medium containing 10% FCS and 20 ng/ml recombinant interleukin 2 (rIL-2) (EuroCetus, France).

Every 3 days the culture fluid was harvested and replaced with fresh medium containing IL-2 and the amount of p24 antigen in the cell-free supernatants was measured by ELISA. Inhibition of infection by human recombinant RANTES (Biodesign International, Kennebunk, ME) was performed by adding 125 nmol final concentration at the time of infection and each time the medium was replaced.

Functional analysis of both CCR5 alleles cloned from ExU2 using an 7Env-mediated cell fusion assay.

A CD4 + HeLa (lacZ) cell clone was generated in HeLa cells carrying a stably integrated lacZ gene under HIV-1 LTR control by stable transfection with a retroviral vector containing CD4 cDNA (P4 HeLa cell clone) (19). Stable surface expression of the HIV-1 LAI envelope glycoprotein in human HeLa cells (HeLa LAI cells) was obtained by transfection of pMA243, an HIV-1 provirus derived from an infectious HIV-1 LAI provirus in which the *gag* and *pol* genes are deleted and the *nef* gene is replaced by the *dhfr* drug resistance gene (20). Stable surface expression of the HIV-1 Ada envelope glycoprotein in HeLa cells (HeLa Ada cells) was obtained by substituting the HIV-1 LAI envelope glycoprotein by the HIV-1

Ada envelope glycoprotein in pMA243 (21). HeLa cells lacking HIV-1 env expression (HeLa Δenv cells) were obtained by transfection of pMA273, an *env* defective HIV-1 provirus (20). CD4 + HeLa (LTR lacZ) cells were transiently transfected with pcDNA3-based constructs permitting the expression of either CCR5 wild type (CCR5wt), m303 (CCR5m303) or Δ32 (CCR5Δ32) cDNA. After 24 h, cells were cocultured with HeLa Ada, HeLa LAI or HeLa Δenv cells. Cell fusion was evaluated after 24 h by measuring β-galactosidase activity in cell lysates and assessed as described in Oberlin et al., Nature, 1996, **382**, 833.

Genetic analysis of CCR5 from two UU individuals, two ExU individuals, and two siblings of ExU2.

The Δ32 fragment was amplified by PCR from PBMC genomic DNA using the pair of primers corresponding to SEQ ID NO:1 and SEQ ID NO:2. PCR was conducted with 1 µg of genomic DNA using 0.2 mM dNTPs, 0.2 µM primers, and 1.25 U of AmpliTAq Gold polymerase (PE Applied Biosystems, Branchburg, NJ) for 35 cycles (94°C, 40 sec ; 60°C, 40 sec ; 72°C, 40 sec) after an initial 10 min denaturation at 94°C. The resulting PCR products were separated on a 2% nusieve agarose gel. Two fragments of 198 bp and 166 bp corresponding to the wild type and the deleted CCR5 alleles were obtained.

The entire CCR5 gene may be amplified :

- either with the pair of primers SEQ ID NO:5 and SEQ ID NO:6; the amplified product may be used for sequencing said gene ;

- or with the pair of primers SEQ ID NO:3 and SEQ ID NO:4; in such a case, the amplified product may be used to cloning and sequencing of the gene.

PCR of genomic and plasmid DNA (200 ng) was conducted as above for 30 cycles (94°C, 1 min30 ; 49°C (with pair of primers SEQ ID NO:5 and SEQ ID NO:6/55°C

(with the pair of primers SEQ ID NO:3 and SEQ ID NO:4), 1 min ; 72°C, 1 min30) after an initial 10 min denaturation at 94°C. The amplified products were cloned in á pCR3 vector using the TA cloning kit (Invitrogen, Leek, Netherlands). Automatic sequencing was performed using the same primers (ESGS, Research and Development department, Evry, France). For analysis of restriction enzyme products of different CCR5 alleles, the entire CCR5 cDNA was amplified using the same set of primers as above, digested with HincII, and the fragments separated on an agarose gel.

The simultaneous detection of the CCR5Δ32 variant and the CCR5m303 variant may be carried out by amplifying the gene with the pair of primers SEQ ID NO:7 and SEQ ID NO:8. PCR was conducted with 1 µg of genomic ADN using 0.2 mM dNTPs, 0.2 µM primers, and 1.25 U of AmpliTAq Gold polymerase (PE Applied Biosystems, Branchburg, NJ) for 35 cycles (94°C, 40 sec; 56,5°C, 1 min; 72°C, 40 sec) after an initial 10 min denaturation at 94°C. The resulting PCR products were separated on a 2% Nusieve agarose gel. For analysis of restriction enzyme, the resulting PCR products are digested with HincII, and the fragments separated on an agarose gel.

The sizes of the obtained products are as follows :

gene :

wt : 460 bp

Δ32 : 428 bp

m303 : 460 bp

after HincII restriction :

wt: 66 and 394 bp

Δ32: 66 and 362 bp

m303: 460 bp.

Infection of PBMC from family members of ExU2 with CCR5-dependent (YU2) and CXCR4-dependent (NL4-3) viruses.

Ficoll purified PBMC were activated as above for 4 days and infected after 1 day with 0.5 ng p24/10^6 cells of either YU2 or NL4-3 HIV molecular clones in RPMI 1640 medium supplemented with 10% FCS, 1 U/ml anti-αIFN (Valbiotech, Paris, France) and 2 µg/ml polybrene (Sigma, France).

RESULTS AND DISCUSSION

ExU1 and ExU2 are two causasian homosexual men who reported multiple incidents of unprotected sexual intercourse during the last ten years. ExU2 reported numerous sexual relationships with multiple partners who succumbed to AIDS and is now the stable partner of an HIV-infected person. No evidence of HIV-1 infection was detected in either ExU individual by standard techniques (HIV-1 ELISA and RNA PCR, Roche), nor was there evidence of clinical or immune alterations as determined by the CD4/CD8 ratio and other parameters of cellular immunity.

The CCR5 genotypes of ExU1 and ExU2 were first determined by PCR of genomic DNA using a pair of primers which allow amplification of the DNA fragment containing the characteristic Δ32 base pair deleted sequence of CCR5Δ32 (2). ExU1 was found to be homozygous for the deleted allele (Δ/Δ), while ExU2 displayed only one deleted allele (wt?/Δ), where wt? indicates that one allele of CCR5 in this individual lacked the 32 base pair deletion and was therefore presumed to be wild type.

There is evidence for an association between CCR5Δ32 homozygoty and resistance to infection (1,2) whereas the same deletion, when present as a heterozygous trait, does not confer resistance (3,11). Cells of

CCR5Δ32 heterozygous individuals (like ExU2) are typically susceptible to infection by CCR5-dependent viruses (1).

To determine whether the presumably exposed but uninfected individuals were resistant to infection, PBMC from both individuals (ExU1 and ExU2) were tested for their ability to be infected by two molecular clones of HIV-1: YU2, which uses CCR5 as a co-receptor (12) and NL4-3, which is strictly CXCR4-dependent (13). PBMC from two uninfected unexposed (UU) individuals who where either homozygous (wt/wt) for the wild type allele (UU1) or heterozygous (wt/Δ) for the deleted allele (UU2) were used as controls.

As expected, PBMC from all four individuals were readily infected by NL4-3 (Fig. 1A-D). In contrast, both ExU1 and ExU2 were completely resistant to infection by YU2 (Fig. 1C and D) while both UU1 and UU2 were susceptible (Fig. 1A and B). Infection of both UU with YU2 could be inhibited by the HIV-suppressive chemokine RANTES, a ligand of CCR5 (Fig. 1A and B).

The resistance of ExU2 PBMC to infection by CCR5-dependent viruses was not restricted to the YU2 molecular clone, as inoculation of PBMC from ExU2 with either two additional CCR5-dependent HIV viruses (JRCSF and BaL) as well as the HIV virus (V164) isolated from the seropositive partner of ExU2, (which, based on susceptibility to neutralization by RANTES is CCR5-dependent) did not result in infection. ExU2 PBMC were equally resistant to the three HIV-1 isolates as well as to the clinical isolate V164 . In contrast, the control UU2 PBMC were infected by all viruses tested. Therefore, despite the fact that ExU2 was apparently heterozygous for CCR5Δ32, he was resistant to infection by CCR5-

dependent HIV-1.

The resistance of ExU2 PBMC to *in vitro* infection by CCR5-dependent viruses, together with his healthy clinical status despite an extensive history of sexual exposure, prompted the search for an alternative mutation within the non-deleted CCR5 allele. Therefore, the entire CCR5 gene of ExU2 was cloned and the nucleotide sequence of both alleles was determined. The presence of the Δ32 deletion in one of the two alleles was confirmed. In addition, a single point mutation (T-->A) at position 303 was found in the non-deleted allele (thereafter called m303). The remainder of the sequence (SEQ ID NO:14) was otherwise identical to the wild type gene (SEQ ID NO:12) (14). The CCR5m303 mutation generates a stop codon which truncates the CCR5 at position 303, resulting in the loss of the transmembrane regions 3-7 and the C-terminal cytoplasmic end of the molecule.

The CCR5m303 mutation together with the Δ32 deletion could account for the incapacity of macrophage-tropic HIV isolates to infect ExU2 PBMC if both proteins expressed by the two alleles were non-functional. To expose this hypothesis, the wild type CCR5 and the two mutant alleles were each transfected into $CD4^+$ human cells and tested for their capacity to generate a functional protein in an HIV envelope-mediated cell fusion assay using $CD4^+$ HeLa cells carrying an integrated HIV long terminal repeat (LTR)-driven reporter gene (*lacZ*). *Env*-mediated cell fusion was assessed by measuring β-galactosidase generated from *lacZ* when HeLa cells transfected with defective provirus but expressing an HIV *env* gene were combined with CD4+, LTR/*lacZ*-containing HeLa cells transfected with nucleic acid encoding wild type CCR5, the CCR5m303 variant or the CCR5Δ32 variant. If

cell fusion via a .$CD4^+$.CCR5.*env* interaction occurs, provirus encoded Tat protein from the *env*-bearing cells can activate the LTR sequence in the CCR5-bearing cells and β-galactosidase expression will occur. The amount of β-galactosidase generated in HeLa $CD4^+$ cell depends on the induction of the HIV-1 LTR by the Tat protein and is an accurate measurement of *Env*-mediated cell fusion.

As shown in Fig. 2, neither the CCR5m303 or the CCR5Δ32 variant were capable of generating a functional, fusion-permissive co-receptor, although wild type CCR5-expressing CD4+ HeLa cells resulted in cell fusion. These results suggest that the resistance to infection of ExU2 was due to the absence of a functional CCR5 co-receptor.

Next, it was tested whether the CCR5m303 mutant was a *de novo* mutation that arose in ExU2 or rather was inherited as a Mendelian trait. CCR5 genetic analysis and PBMC infection assays were performed in two immediate relatives of ExU2, the father (2.F) and his sister (2.S). When tested for the presence of the Δ32 mutation, it was found that both father and sister were heterozygous for this deletion (Fig. 3B).

Next, we took advantage of the presence of a HincII restriction site at position 303 which is lost by the T--> A mutation, but present in the wild type CCR5 gene (Fig. 3A) to screen genomic DNA from all individuals in this study. The entire CCR5 genomic sequences corresponding to ExU2, 2.S, 2.F, ExU1, UU1 and UU2 were amplified by PCR. Figure 3C lanes 1-6 shows the uncleaved amplified DNA prepared from these six individuals, separated on a 1% agarose gel; all alleles showed similar patterns of migration. However, digestion with the restriction enzyme HincII generated different restriction

patterns in the CCR5 alleles. Two distincts fragments of 0.7 and 0.4 kb were obtained from the genomic DNA of subjects who were either wild type (UU1), homozygous (ExU1) or heterozygous (UU2) for the Δ32 allele (Fig. 3D, lanes 4, 5 and 6). In contrast, the pattern of ExU2, which carries the CCR5m303 mutation on one allele, displayed an additional band migrating with an apparent size of 1,1 kb (Fig. 3D, lane 1).

To confirm these findings, we analyzed cDNAs corresponding to CCR5 wild type, CCR5Δ32 and CCR5m303 alleles which had been characterized by nucleotide sequencing and cloned in a pCR3.1 plasmid. All of the cDNA products displayed the same restriction enzyme characteristics and migration pattern as those observed for CCR5 alleles directly amplified from genomic DNA (Fig. 3C, lanes 7, 8 and 9). HincII digestion of the three cloned genes generated two distinct fragments of 0,7 and 0,3 kb only within the wild type and CCR5Δ32 clones (Fig. 3D, lanes 7 and 8). In contrast, the CCR5m303 clone was not digested and displayed the original migration pattern of the entire CCR5 gene (1.1 kb) (Fig. 3D, lane 9). Upon digestion with HincII, genomic DNA from 2.S but not 2.F, displayed a migration pattern identical to that of ExU2 (Fig. 3D, lanes 2 and 3), suggesting that the CCR5m303 mutation was inherited by both siblings from the mother as a single mendelian trait. The presence of the Δ32 and m303 alleles on cDNA from 2.S was confirmed by automatic nucleotide sequencing of the corresponding clones.

Since both ExU2 and 2.S carry the same CCR5 genetic trait, i.e., a combination of the CCR5m303 mutant and CCR5Δ32 deleted alleles, we compared their *in vitro* resistance to infection by a CCR5-dependent virus (YU2).

In addition, we tested PBMC from 2.F, which carry only the CCR5Δ32 deleted allele along with the wild type one. As would be expected based upon the lack of functionality of the two mutant alleles, 2.S PBMC displayed the same degree of resistance to infection as ExU2 PBMC whereas the 2.F PBMC were fully susceptible to infection with YU2 (Fig. 3A). All three individuals were susceptible to *in vitro* infection with NL4-3, the CXCR4-dependent virus (Fig. 3B).

Figure 5A-B.

Figure 5A Hela P4 cells were cotransfected with different plasmid containing either CCR5 wild type (wt), CCR5Δ32 (Δ) or CCR5m303 (m), with DNA ratio (1/1; 3μg/3μg). After 48h, an *Env*-mediated cell fusion assay was conducted with HeLa Ada cells. A ß galactosidase assay was realised and the results expressed by the number of blue cells. The results illustrated in said Figure 5A suggest that the negative interference of the mutant CCR5m303 is more effective than the CCR5Δ32 mutant.

Figure 5B: U373 cells were cotransfected with different plasmid containing either CCR5 wild type (wt), CCR5Δ32 (Δ) or CCR5m303 (m), with various DNA ratio (1/1, 2μg/2μg; 1/3, 2μg/6μg; 3/1,6μg/2μg). After 48h, a cell fusion assay was conducted with HeLa Ada cells. 24 hours later a CPRG lysis test assay was conducted and the results are expressed with an OD value at 540nm. The experience show that with CCR5m303 whatever the ratio the same negative interference was observed. On the contrary, a higher amount of wild type plasmid could suppressed the negative interference of the CCR5Δ32. It seems that the negative interference observed with the CCR5Δ32 is dose dependent which is not the case for the CCR5m303.

REFERENCES

1. M. Samson et al., *Nature* 382, 722 (1996)
2. R. Lju et al., *Cell* 86, 367 (1996)
3. M. Dean et al., *Science* 273, 1856 (1996)
4. T. Dragic et al., *Nature* 381, 667 (1996)
5. H. Deng et al., *Nature* 381, 661 (1996)
6. G. Alkhatib et al., *Science* 272, 1955 (1996)
7. H. Choe et al., *Cell* 85, 1135 (1996)
8. T. Zhu et al., *Science* 261, 1179 (1993)
9. B.J. Doranz et al., *Cell* 85, 1149 (1996)
10. Y. Feng, C.C. Broder, P.E. Kennedy, E.A. Berger, *Science* 272, 872 (1996)
11. Y. Huang et al., *Nature Med.* 2, 1240 (1996)
12. Y. Li et al., *J. Virol.* 66, 6597 (1992)
13. L. Zhang, Y. Huang, Y. He, Y. Cao, D.D. Ho, *Nature* 383, 768 (1996)
14. M. Samson, O/ Labbe, C. Mollereau, G. Vassart, M. Parmentier, *Biochemistry* 35, 3362 (1996)
15. G. Shearer and M. Clerici, *Immun Today* 17, 21 (1996)
16. W.A. Paxton et al., *Nature Med.* 2, 412 (1996)
17. R. Biti, R. French, J. Young, B/ Bennetts, G. Stewart, *Nature Med.* 3, 252 (1997)
18. J.P. Moore, *Science* 276, 51 (1997)
19. F. Clavel and P. Chameau, *J. Virol.* 68, 1179 (1994)
20. T. Dragic, P. Chameau, F. Clavel, M. Alizon, *J. Virol.* 68, 4794 (1992)
21. O. Pleskoff et al., *Science 276,* 1874-1898 (1997),
22. E. Oberlin et al., *Nature* 382, 833 (1996).

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: FONDATION MONDIALE RECHERCHE ET PREVENTION SIDA
(B) STREET: 1 RUE MIOLLIS
(C) CITY: PARIS
(E) COUNTRY: FRANCE
(F) POSTAL CODE (ZIP): 75732 CEDEX 15

(ii) TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS CO-RECEPTOR VARIANTS ASSOCIATED WITH RESISTANCE TO VIRUS INFECTION.

(iii) NUMBER OF SEQUENCES: 17

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GTCTTCATTA CACCTGCAGC TC 22

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GTGAAGATAA GCCTCACAGC C 21

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

CCCAAGCTTA TGGATTATCA AGTG 24

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GCTCTAGATC ACAAGCCCAC AGA 23

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GGGCAACTAA ATACAT 16

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GCACAACTCT GACTGG 16

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

TCCTTCTTAC TGTCCCCTTC TGG 23

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

CCTGTGCCTC TTCTTCTCAT TTC 23

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

AATACAATGT GTCAACTCTT G 21

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

AATACAATGT GACAACTCTT G 21

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 461 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

TCCTTCTTAC TGTCCCCTTC TGGGCTCACT ATGCTGCCGC CCAGTGGGAC TTTGGAAATA 60

CAATGTGACA ACTCTTGACA GGGCTCTATT TTATAGGCTT CTTCTCTGGA ATCTTCTTCA 120

TCATCCTCCT GACAATCGAT AGGTACCTGG CTGTCGTCCA TGCTGTGTTT GCTTTAAAAG 180

CCAGGACGGT CACCTTTGGG GTGGTGACAA GTGTGATCAC TTGGGTGGTG GCTGTGTTTG 240

CGTCTCTCCC AGGAATCATC TTTACCAGAT CTCAAAAAGA AGGTCTTCAT TACACCTGCA 300

GCTCTCATTT TCCATACAGT CAGTATCAAT TCTGGAAGAA TTTCCAGACA TTAAAGATAG 360

TCATCTTGGG GCTGGTCCTG CCGCTGCTTG TCATGGTCAT CTGCTACTCG GGAATCCTAA 420

AAACTCTGCT TCGGTGTCGA AATGAGAAGA AGAGGCACAG G 461

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1376 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:240..1298

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

```
GAATTCCCCC AACAGAGCCA AGCTCTCCAT CTAGTGGACA GGGAAGCTAG CAGCAAACCT    60

TCCCTTCACT ACAAAACTTC ATTGCTTGGC CAAAAAGAGA GTTAATTCAA TGTAGACATC   120

TATGTAGGCA ATTAAAAACC TATTGATGTA TAAAACAGTT TGCATTCATG GAGGGCAACT   180

AAATACATTC TAGGACTTTA TAAAAGATCA CTTTTTATTT ATGCACAGGG TGGAACAAG    239

ATG GAT TAT CAA GTG TCA AGT CCA ATC TAT GAC ATC AAT TAT TAT ACA      287
Met Asp Tyr Gln Val Ser Ser Pro Ile Tyr Asp Ile Asn Tyr Tyr Thr
  1               5                  10                  15

TCG GAG CCC TGC CAA AAA ATC AAT GTG AAG CAA ATC GCA GCC CGC CTC      335
Ser Glu Pro Cys Gln Lys Ile Asn Val Lys Gln Ile Ala Ala Arg Leu
             20                  25                  30

CTG CCT CCG CTC TAC TCA CTG GTG TTC ATC TTT GGT TTT GTG GGC AAC      383
Leu Pro Pro Leu Tyr Ser Leu Val Phe Ile Phe Gly Phe Val Gly Asn
         35                  40                  45

ATG CTG GTC ATC CTC ATC CTG ATA AAC TGC AAA AGG CTG AAG AGC ATG      431
Met Leu Val Ile Leu Ile Leu Ile Asn Cys Lys Arg Leu Lys Ser Met
     50                  55                  60

ACT GAC ATC TAC CTG CTC AAC CTG GCC ATC TCT GAC CTG TTT TTC CTT      479
Thr Asp Ile Tyr Leu Leu Asn Leu Ala Ile Ser Asp Leu Phe Phe Leu
 65                  70                  75                  80

CTT ACT GTC CCC TTC TGG GCT CAC TAT GCT GCC GCC CAG TGG GAC TTT      527
Leu Thr Val Pro Phe Trp Ala His Tyr Ala Ala Ala Gln Trp Asp Phe
                 85                  90                  95

GGA AAT ACA ATG XXX CAA CTC TTG ACA GGG CTC TAT TTT ATA GGC TTC      575
Gly Asn Thr Met YYY Gln Leu Leu Thr Gly Leu Tyr Phe Ile Gly Phe
            100                 105                 110
```

```
TTC TCT GGA ATC TTC TTC ATC ATC CTC CTG ACA ATC GAT AGG TAC CTG      623
Phe Ser Gly Ile Phe Phe Ile Ile Leu Leu Thr Ile Asp Arg Tyr Leu
        115                 120                 125

GCT GTC GTC CAT GCT GTG TTT GCT TTA AAA GCC AGG ACG GTC ACC TTT      671
Ala Val Val His Ala Val Phe Ala Leu Lys Ala Arg Thr Val Thr Phe
    130                 135                 140

GGG GTG GTG ACA AGT GTG ATC ACT TGG GTG GTG GCT GTG TTT GCG TCT      719
Gly Val Val Thr Ser Val Ile Thr Trp Val Val Ala Val Phe Ala Ser
145                 150                 155                 160

CTC CCA GGA ATC ATC TTT ACC AGA TCT CAA AAA GAA GGT CTT CAT TAC      767
Leu Pro Gly Ile Ile Phe Thr Arg Ser Gln Lys Glu Gly Leu His Tyr
                165                 170                 175

ACC TGC AGC TCT CAT TTT CCA TAC AGT CAG TAT CAA TTC TGG AAG AAT      815
Thr Cys Ser Ser His Phe Pro Tyr Ser Gln Tyr Gln Phe Trp Lys Asn
            180                 185                 190

TTC CAG ACA TTA AAG ATA GTC ATC TTG GGG CTG GTC CTG CCG CTG CTT      863
Phe Gln Thr Leu Lys Ile Val Ile Leu Gly Leu Val Leu Pro Leu Leu
        195                 200                 205

GTC ATG GTC ATC TGC TAC TCG GGA ATC CTA AAA ACT CTG CTT CGG TGT      911
Val Met Val Ile Cys Tyr Ser Gly Ile Leu Lys Thr Leu Leu Arg Cys
    210                 215                 220

CGA AAT GAG AAG AAG AGG CAC AGG GCT GTG AGG CTT ATC TTC ACC ATC      959
Arg Asn Glu Lys Lys Arg His Arg Ala Val Arg Leu Ile Phe Thr Ile
225                 230                 235                 240

ATG ATT GTT TAT TTT CTC TTC TGG GCT CCC TAC AAC ATT GTC CTT CTC     1007
Met Ile Val Tyr Phe Leu Phe Trp Ala Pro Tyr Asn Ile Val Leu Leu
                245                 250                 255

CTG AAC ACC TTC CAG GAA TTC TTT GGC CTG AAT AAT TGC AGT AGC TCT     1055
Leu Asn Thr Phe Gln Glu Phe Phe Gly Leu Asn Asn Cys Ser Ser Ser
            260                 265                 270

AAC AGG TTG GAC CAA GCT ATG CAG GTG ACA GAG ACT CTT GGG ATG ACG     1103
Asn Arg Leu Asp Gln Ala Met Gln Val Thr Glu Thr Leu Gly Met Thr
        275                 280                 285

CAC TGC TGC ATC AAC CCC ATC ATC TAT GCC TTT GTC GGG GAG AAG TTC     1151
His Cys Cys Ile Asn Pro Ile Ile Tyr Ala Phe Val Gly Glu Lys Phe
    290                 295                 300

AGA AAC TAC CTC TTA GTC TTC TTC CAA AAG CAC ATT GCC AAA CGC TTC     1199
Arg Asn Tyr Leu Leu Val Phe Phe Gln Lys His Ile Ala Lys Arg Phe
305                 310                 315                 320
```

```
TGC AAA TGC TGT TCT ATT TTC CAG CAA GAG GCT CCC GAG CGA GCA AGC     1247
Cys Lys Cys Cys Ser Ile Phe Gln Gln Glu Ala Pro Glu Arg Ala Ser
                325                 330                 335

TCA GTT TAC ACC CGA TCC ACT GGG GAG CAG GAA ATA TCT GTG GGC TTG     1295
Ser Val Tyr Thr Arg Ser Thr Gly Glu Gln Glu Ile Ser Val Gly Leu
            340                 345                 350

TGA CACGGACTCA AGTGGGCTGG TGACCCAGTC AGAGTTGTGC ACATGGCTTA          1348
 *

GTTTTCATAC ACAGCCTGGG CTGGGGGT                                      1376
```

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 353 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

```
Met Asp Tyr Gln Val Ser Ser Pro Ile Tyr Asp Ile Asn Tyr Tyr Thr
  1               5                  10                  15

Ser Glu Pro Cys Gln Lys Ile Asn Val Lys Gln Ile Ala Ala Arg Leu
            20                  25                  30

Leu Pro Pro Leu Tyr Ser Leu Val Phe Ile Phe Gly Phe Val Gly Asn
         35                  40                  45

Met Leu Val Ile Leu Ile Leu Ile Asn Cys Lys Arg Leu Lys Ser Met
     50                  55                  60

Thr Asp Ile Tyr Leu Leu Asn Leu Ala Ile Ser Asp Leu Phe Phe Leu
 65                  70                  75                  80

Leu Thr Val Pro Phe Trp Ala His Tyr Ala Ala Ala Gln Trp Asp Phe
                85                  90                  95

Gly Asn Thr Met Cys Gln Leu Leu Thr Gly Leu Tyr Phe Ile Gly Phe
            100                 105                 110

Phe Ser Gly Ile Phe Phe Ile Ile Leu Leu Thr Ile Asp Arg Tyr Leu
        115                 120                 125

Ala Val Val His Ala Val Phe Ala Leu Lys Ala Arg Thr Val Thr Phe
    130                 135                 140

Gly Val Val Thr Ser Val Ile Thr Trp Val Val Ala Val Phe Ala Ser
145                 150                 155                 160
```

```
Leu Pro Gly Ile Ile Phe Thr Arg Ser Gln Lys Glu Gly Leu His Tyr
                165                 170                 175

Thr Cys Ser Ser His Phe Pro Tyr Ser Gln Tyr Gln Phe Trp Lys Asn
            180                 185                 190

Phe Gln Thr Leu Lys Ile Val Ile Leu Gly Leu Val Leu Pro Leu Leu
        195                 200                 205

Val Met Val Ile Cys Tyr Ser Gly Ile Leu Lys Thr Leu Leu Arg Cys
    210                 215                 220

Arg Asn Glu Lys Lys Arg His Arg Ala Val Arg Leu Ile Phe Thr Ile
225                 230                 235                 240

Met Ile Val Tyr Phe Leu Phe Trp Ala Pro Tyr Asn Ile Val Leu Leu
                245                 250                 255

Leu Asn Thr Phe Gln Glu Phe Phe Gly Leu Asn Asn Cys Ser Ser Ser
            260                 265                 270

Asn Arg Leu Asp Gln Ala Met Gln Val Thr Glu Thr Leu Gly Met Thr
        275                 280                 285

His Cys Cys Ile Asn Pro Ile Ile Tyr Ala Phe Val Gly Glu Lys Phe
    290                 295                 300

Arg Asn Tyr Leu Leu Val Phe Phe Gln Lys His Ile Ala Lys Arg Phe
305                 310                 315                 320

Cys Lys Cys Cys Ser Ile Phe Gln Gln Glu Ala Pro Glu Arg Ala Ser
                325                 330                 335

Ser Val Tyr Thr Arg Ser Thr Gly Glu Gln Glu Ile Ser Val Gly Leu
            340                 345                 350
```

*

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1071 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:7..309

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

```
AAGCTT ATG GAT TAT CAA GTG TCA AGT CCA ATC TAT GAC ATC AAT TAT        48
       Met Asp Tyr Gln Val Ser Ser Pro Ile Tyr Asp Ile Asn Tyr
           355                 360                 365

TAT ACA TCG GAG CCC TGC CAA AAA ATC AAT GTG AAG CAA ATC GCA GCC      96
Tyr Thr Ser Glu Pro Cys Gln Lys Ile Asn Val Lys Gln Ile Ala Ala
        370                 375                 380

CGC CTC CTG CCT CCG CTC TAC TCA CTG GTG TTC ATC TTT GGT TTT GTG     144
Arg Leu Leu Pro Pro Leu Tyr Ser Leu Val Phe Ile Phe Gly Phe Val
    385                 390                 395

GGC AAC ATG CTG GTC ATC CTC ATC CTG ATA AAC TGC AAA AGG CTG AAG     192
Gly Asn Met Leu Val Ile Leu Ile Leu Ile Asn Cys Lys Arg Leu Lys
400                 405                 410                 415

AGC ATG ACT GAC ATC TAC CTG CTC AAC CTG GCC ATC TCT GAC CTG TTT     240
Ser Met Thr Asp Ile Tyr Leu Leu Asn Leu Ala Ile Ser Asp Leu Phe
                420                 425                 430

TTC CTT CTT ACT GTC CCC TTC TGG GCT CAC TAT GCT GCC GCC CAG TGG     288
Phe Leu Leu Thr Val Pro Phe Trp Ala His Tyr Ala Ala Ala Gln Trp
            435                 440                 445

GAC TTT GGA AAT ACA ATG TGA CAACTCTTGA CAGGGCTCTA TTTTATAGGC        339
Asp Phe Gly Asn Thr Met  *
        450

TTCTTCTCTG GAATCTTCTT CATCATCCTC CTGACAATCG ATAGGTACCT GGCTGTCGTC   399

CATGCTGTGT TTGCTTTAAA AGCCAGGACG GTCACCTTTG GGGTGGTGAC AAGTGTGATC   459

ACTTGGGTGG TGGCTGTGTT TGCGTCTCTC CCAGGAATCA TCTTTACCAG ATCTCAAAAA   519

GAAGGTCTTC ATTACACCTG CAGCTCTCAT TTTCCATACA GTCAGTATCA ATTCTGGAAG   579

AATTTCCAGA CATTAAAGAT AGTCATCTTG GGGCTGGTCC TGCCGCTGCT TGTCATGGTC   639

ATCTGCTACT CGGGAATCCT AAAAACTCTG CTTCGGTGTC GAAATGAGAA GAAGAGGCAC   699

AGGGCTGTGA GGCTTATCTT CACCATCATG ATTGTTTATT TTCTCTTCTG GGCTCCCTAC   759

AACATTGTCC TTCTCCTGAA CACCTTCCAG GAATTCTTTG GCCTGAATAA TTGCAGTAGC   819

TCTAACAGGT TGGACCAAGC TATGCAGGTG ACAGAGACTC TTGGGATGAC GCACTGCTGC   879

ATCAACCCCA TCATCTATGC CTTTGTCGGG GAGAAGTTCA GAAACTACCT CTTAGTCTTC   939

TTCCAAAAGC ACATTGCCAA ACGCTTCTGC AAATGCTGTT CTATTTTCCA GCAAGAGGCT   999
```

CCCGAGCGAG CAAGCTCAGT TTACACCCGA TCCACTGGGG AGCAGGAAAT ATCTGTGGGC 1059

TTGTGATCTA GA 1071

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 101 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

Met Asp Tyr Gln Val Ser Ser Pro Ile Tyr Asp Ile Asn Tyr Tyr Thr
1 5 10 15

Ser Glu Pro Cys Gln Lys Ile Asn Val Lys Gln Ile Ala Ala Arg Leu
20 25 30

Leu Pro Pro Leu Tyr Ser Leu Val Phe Ile Phe Gly Phe Val Gly Asn
35 40 45

Met Leu Val Ile Leu Ile Leu Ile Asn Cys Lys Arg Leu Lys Ser Met
50 55 60

Thr Asp Ile Tyr Leu Leu Asn Leu Ala Ile Ser Asp Leu Phe Phe Leu
65 70 75 80

Leu Thr Val Pro Phe Trp Ala His Tyr Ala Ala Ala Gln Trp Asp Phe
85 90 95

Gly Asn Thr Met *
100

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1344 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:240..887

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

```
GAATTCCCCC AACAGAGCCA AGCTCTCCAT CTAGTGGACA GGGAAGCTAG CAGCAAACCT     60

TCCCTTCACT ACAAAACTTC ATTGCTTGGC CAAAAAGAGA GTTAATTCAA TGTAGACATC    120

TATGTAGGCA ATTAAAAACC TATTGATGTA TAAAACAGTT TGCATTCATG GAGGGCAACT    180

AAATACATTC TAGGACTTTA TAAAAGATCA CTTTTTATTT ATGCACAGGG TGGAACAAG     239

ATG GAT TAT CAA GTG TCA AGT CCA ATC TAT GAC ATC AAT TAT TAT ACA       287
Met Asp Tyr Gln Val Ser Ser Pro Ile Tyr Asp Ile Asn Tyr Tyr Thr
            105                 110                 115

TCG GAG CCC TGC CAA AAA ATC AAT GTG AAG CAA ATC GCA GCC CGC CTC      335
Ser Glu Pro Cys Gln Lys Ile Asn Val Lys Gln Ile Ala Ala Arg Leu
        120                 125                 130

CTG CCT CCG CTC TAC TCA CTG GTG TTC ATC TTT GGT TTT GTG GGC AAC      383
Leu Pro Pro Leu Tyr Ser Leu Val Phe Ile Phe Gly Phe Val Gly Asn
    135                 140                 145

ATG CTG GTC ATC CTC ATC CTG ATA AAC TGC AAA AGG CTG AAG AGC ATG      431
Met Leu Val Ile Leu Ile Leu Ile Asn Cys Lys Arg Leu Lys Ser Met
150                 155                 160                 165

ACT GAC ATC TAC CTG CTC AAC CTG GCC ATC TCT GAC CTG TTT TTC CTT      479
Thr Asp Ile Tyr Leu Leu Asn Leu Ala Ile Ser Asp Leu Phe Phe Leu
                170                 175                 180

CTT ACT GTC CCC TTC TGG GCT CAC TAT GCT GCC GCC CAG TGG GAC TTT      527
Leu Thr Val Pro Phe Trp Ala His Tyr Ala Ala Ala Gln Trp Asp Phe
            185                 190                 195

GGA AAT ACA ATG TGT CAA CTC TTG ACA GGG CTC TAT TTT ATA GGC TTC      575
Gly Asn Thr Met Cys Gln Leu Leu Thr Gly Leu Tyr Phe Ile Gly Phe
        200                 205                 210

TTC TCT GGA ATC TTC TTC ATC ATC CTC CTG ACA ATC GAT AGG TAC CTG      623
Phe Ser Gly Ile Phe Phe Ile Ile Leu Leu Thr Ile Asp Arg Tyr Leu
    215                 220                 225

GCT GTC GTC CAT GCT GTG TTT GCT TTA AAA GCC AGG ACG GTC ACC TTT      671
Ala Val Val His Ala Val Phe Ala Leu Lys Ala Arg Thr Val Thr Phe
230                 235                 240                 245

GGG GTG GTG ACA AGT GTG ATC ACT TGG GTG GTG GCT GTG TTT GCG TCT      719
Gly Val Val Thr Ser Val Ile Thr Trp Val Val Ala Val Phe Ala Ser
                250                 255                 260

CTC CCA GGA ATC ATC TTT ACC AGA TCT CAA AAA GAA GGT CTT CAT TAC      767
Leu Pro Gly Ile Ile Phe Thr Arg Ser Gln Lys Glu Gly Leu His Tyr
            265                 270                 275
```

```
ACC TGC AGC TCT CAT TTT CCA TAC ATT AAA GAT AGT CAT CTT GGG GCT      815
Thr Cys Ser Ser His Phe Pro Tyr Ile Lys Asp Ser His Leu Gly Ala
        280                 285                 290

GGT CCT GCC GCT GCT TGT CAT GGT CAT CTG CTA CTC GGG AAT CCT AAA      863
Gly Pro Ala Ala Ala Cys His Gly His Leu Leu Leu Gly Asn Pro Lys
    295                 300                 305

AAC TCT GCT TCG GTG TCG AAA TGA GAAGAAGAGG CACAGGGCTG TGAGGCTTAT     917
Asn Ser Ala Ser Val Ser Lys  *
310                 315

CTTCACCATC ATGATTGTTT ATTTTCTCTT CTGGGCTCCC TACAACATTG TCCTTCTCCT    977

GAACACCTTC CAGGAATTCT TTGGCCTGAA TAATTGCAGT AGCTCTAACA GGTTGGACCA   1037

AGCTATGCAG GTGACAGAGA CTCTTGGGAT GACGCACTGC TGCATCAACC CCATCATCTA   1097

TGCCTTTGTC GGGGAGAAGT TCAGAAACTA CCTCTTAGTC TTCTTCCAAA AGCACATTGC   1157

CAAACGCTTC TGCAAATGCT GTTCTATTTT CCAGCAAGAG GCTCCCGAGC GAGCAAGCTC   1217

AGTTTACACC CGATCCACTG GGGAGCAGGA AATATCTGTG GGCTTGTGAC ACGGACTCAA   1277

GTGGGCTGGT GACCCAGTCA GAGTTGTGCA CATGGCTTAG TTTTCATACA CAGCCTGGGC   1337

TGGGGGT                                                             1344
```

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 216 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

```
Met Asp Tyr Gln Val Ser Ser Pro Ile Tyr Asp Ile Asn Tyr Tyr Thr
 1               5                  10                  15

Ser Glu Pro Cys Gln Lys Ile Asn Val Lys Gln Ile Ala Ala Arg Leu
            20                  25                  30

Leu Pro Pro Leu Tyr Ser Leu Val Phe Ile Phe Gly Phe Val Gly Asn
        35                  40                  45

Met Leu Val Ile Leu Ile Leu Ile Asn Cys Lys Arg Leu Lys Ser Met
    50                  55                  60

Thr Asp Ile Tyr Leu Leu Asn Leu Ala Ile Ser Asp Leu Phe Phe Leu
65                  70                  75                  80
```

```
Leu Thr Val Pro Phe Trp Ala His Tyr Ala Ala Ala Gln Trp Asp Phe
                  85                  90                  95

Gly Asn Thr Met Cys Gln Leu Leu Thr Gly Leu Tyr Phe Ile Gly Phe
             100                 105                 110

Phe Ser Gly Ile Phe Phe Ile Ile Leu Leu Thr Ile Asp Arg Tyr Leu
         115                 120                 125

Ala Val Val His Ala Val Phe Ala Leu Lys Ala Arg Thr Val Thr Phe
     130                 135                 140

Gly Val Val Thr Ser Val Ile Thr Trp Val Val Ala Val Phe Ala Ser
145                 150                 155                 160

Leu Pro Gly Ile Ile Phe Thr Arg Ser Gln Lys Glu Gly Leu His Tyr
                 165                 170                 175

Thr Cys Ser Ser His Phe Pro Tyr Ile Lys Asp Ser His Leu Gly Ala
             180                 185                 190

Gly Pro Ala Ala Ala Cys His Gly His Leu Leu Leu Gly Asn Pro Lys
         195                 200                 205

Asn Ser Ala Ser Val Ser Lys *
     210                 215
```

Claims

1°) A purified CCR5 variant protein which comprises the first two transmembrane domains of wild type CCR5 but which lacks transmembrane domains 3-7.

2°) The purified CCR5 protein variant of claim 1, which is the CCR5 variant CCR5m303 having the amino acid sequence set forth in SEQ ID NO:15.

3°) The purified CCR5 variant protein of claim 2, which is comprised together with heterologous amino acid sequence in a fusion protein.

4°) A purified and isolated nucleic acid sequence encoding the protein of claim 1 or of claim 2.

5°) The nucleic acid of claim 4, which has the sequence set forth in SEQ ID NO:12 where XXX is a stop codon or in SEQ ID NO:14.

6°) Isolated DNA primers for the amplification of the nucleic acid of claim 4, selected from the group consisting of nucleotide fragments, which comprise between 15 and 100 nucleotides of a sequence upstream or downstream nucleotide 540, with reference to SEQ ID NO: 12.

7°) Isolated DNA primers for the amplification of the nucleic acid of claim 4, selected from the group consisting of :

SEQ ID NO:1 and SEQ ID NO:2

SEQ ID NO:3 and SEQ ID NO:4

SEQ ID NO:5 and SEQ ID NO:6

SEQ ID NO:7 and SEQ ID NO:8

8°) Isolated DNA probes for the detection of the nucleic acid of claim 4, selected from the group consisting of SEQ ID NO:10 and SEQ ID NO:11.

9°) Method of detecting and identifying the presence of at least a CCR5 variant in a biological sample, comprising :

- bringing into contact the biological sample with a pair of primers according to claim 6,

- amplifying the CCR5 variant, eventually present in said biological sample

- demonstrating the amplification of the DNA fragment corresponding to the fragment flanked by the primers, for example by gel electrophoresis,

- optionally verifying the sequence of the amplified fragment, for example by specific probe hybridization, sequencing or restriction site analysis.

10°) Method according to claim 9, wherein said pair of primers are selected from the group consisting of:

SEQ ID NO:1 and SEQ ID NO:2

SEQ ID NO:3 and SEQ ID NO:4

SEQ ID NO:5 and SEQ ID NO:6

SEQ ID NO:7 and SEQ ID NO:8,

11°) Method according to claim 9, wherein said hybridization probes are selected in the group consisting of SEQ ID NO :10and SEQ ID NO :11.

12°) A kit ready for use for detecting at least a CCR5 variant in a biological sample, comprising :

DNA primers as claimed in claim 6 or in claim 7,

reagents for extracting the DNA of said biological sample,

reagents for amplifying DNA comprising polymerization enzymes,

optionally, at least one oligonucleotide probe which specifically hybridizes to the amplified DNA, as claimed in claim 8.

13°) A cell into which the nucleic acid of claim 4 or claim 5 has been introduced.

14°) The nucleic acid of claim 4 or claim 5 as

contained in a vector.

15°) A method of determining whether an individual is resistant to infection by a macrophage-tropic strain of human immunodeficiency virus type 1, comprising determining whether a CCR5 variant is present in a sample collected from the individual, wherein the CCR5 variant comprises the first two membrane domains of wild type CCR5 but lacks transmembrane domains 3-7, wherein the presence of such a variant bears a positive correlation with a resistance to the infection.

16°) The method of claim 15, wherein the CCR5 is CCR5m303.

17°) A method for determining the likelihood that an individual will be infected by a macrophage-tropic strain of human immunodeficiency virus type 1, comprising determining whether a CCR5 variant is present in a sample collected from the individual, wherein the CCR5 variant comprises the first two transmembrane domains of wild type CCR5 but lacks transmembrane domains 3-7, wherein the presence of such a variant bears a positive correlation with a lower probability of infection.

18°) A method of evaluating the prognosis of an individual infected with a macrophage-tropic strain of human immunodeficiency virus type 1, comprising determining whether a CCR5 variant is present in a sample collected from the individual, wherein the CCR5 variant comprises the first two transmembrane domains of wild type CCR5 but lacks transmembrane domains 3-7, wherein the presence of such a variant bears a positive correlation with a slower progression of the disease.

19°) A method of determining whether an individual is resistant to infection by a macrophage-tropic strain of human immunodeficiency virus type 1, comprising characterizing both alleles of the individual, wherein a

determination that both CCR5 alleles of an individual encode a CCR5 variant which either (i) comprises the first two transmembrane domains of wild type CCR5 but which lacks transmembrane domains 3-7 or (ii) is functionally defective (e.g.) cannot be expressed at the cell surface) has a positive correlation with a resistance of the individual to infection.

20°) The method of claims 15 to 19, wherein the presence of the CCR5 variant is determined by identifying a nucleic acid which encodes said variant.

21°) The method of claims 15 to 19, wherein said identifying step is performed according to the method of claims 9 to 11.

22°) The method of claim 19, wherein one CCR5 allele is a CCR5m303 variant and wherein the other CCR5 allele is a CCR5Δ32 variant.

23°) A method of determining whether an individual is less likely to be infected by a macrophage-tropic strain of human immunodeficiency virus type 1, comprising characterizing both CCR5 alleles of the individual, wherein a determination that one CCR5 allele is a wild type CCR5 allele and the other allele is a CCR5 variant which comprises the first two transmembrane domains of wild type CCR5 but which lacks transmembrane domains 3-7 has a positive correlation with a lower probability of infection relative to homozygous CCR5 wild type individuals.

24°) A method of evaluating the prognosis of an individual infected with a macrophage-tropic strain of human immunodeficiency virus type 1, comprising characterizing both CCR5 alleles of the individual, wherein a determination that one CCR5 allele is a wild type CCR5 allele and the other is a CCR5 variant which comprises the first two transmembrane domains of the wild type CCR5

but which lacks transmembrane domains 3-7 has a positive correlation with a slower progression of disease.

25°) Composition characterized in that it comprises a nucleic acid according to claim 4 or 5, for the preparation of a drug for inhibiting CCR5-mediated HIV infection.

26°) Method of detecting a CCR5 variant protein according to claims 1 to 3, wherein it comprises:

- contacting a biological sample with an antibody directed against said CCR5 variant protein, and
- detecting the complex CCR5 variant protein - antibodies by an appropriate means.

27°) Antibodies, characterized in that they are directed against a CCR5 variant protein according to claims 1 to 3.

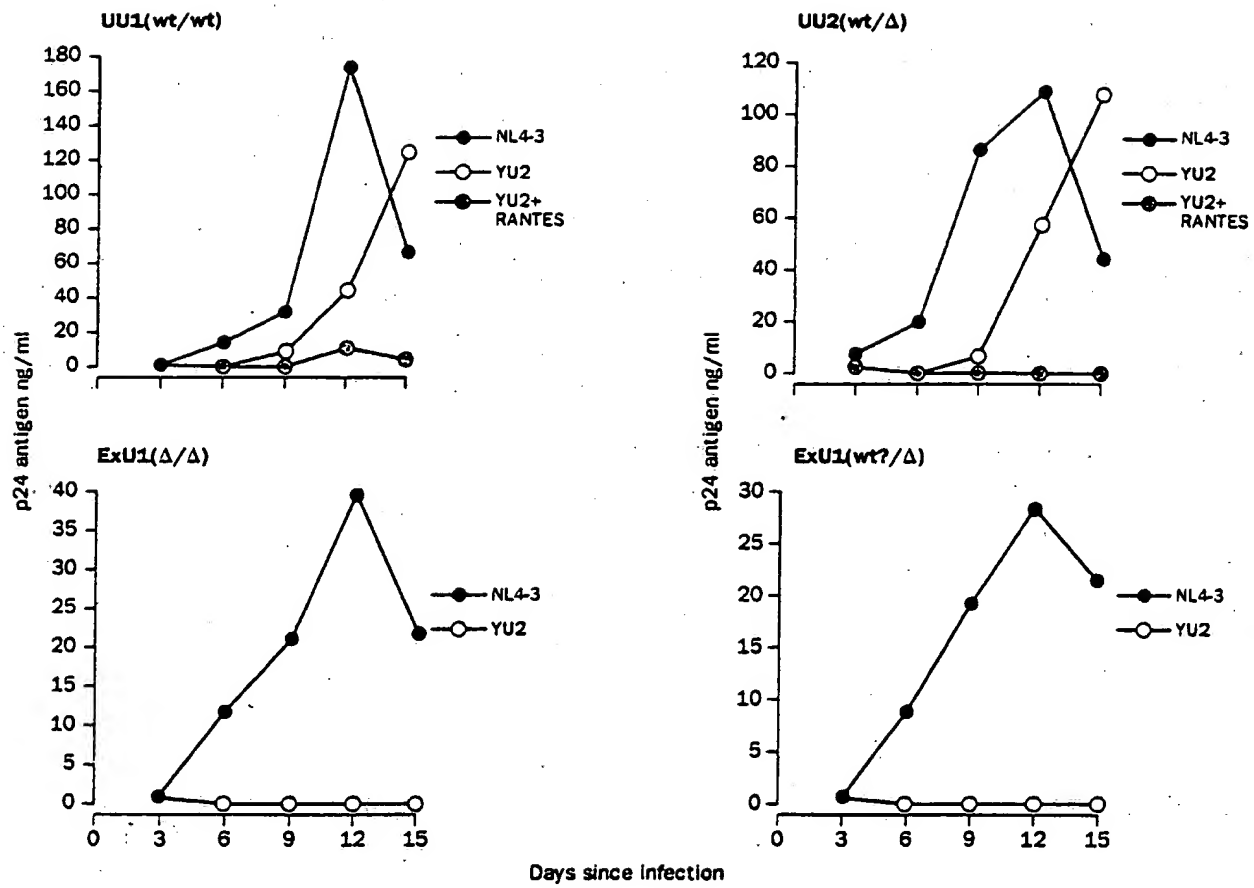

UU1(wt/wt)
UU2(wt/Δ)
ExU1(Δ/Δ)
ExU1(wt?/Δ)
NL4-3
YU2
YU2+ RANTES
p24 antigen ng/ml
Days since infection

FIGURE 1

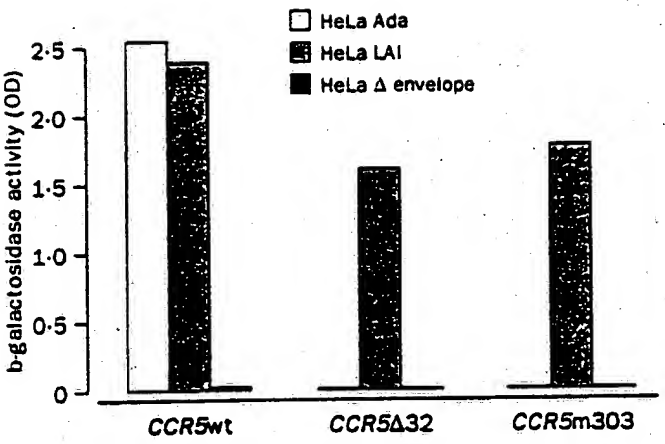
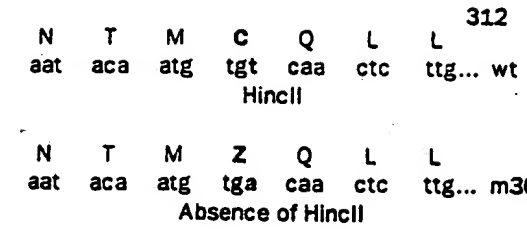
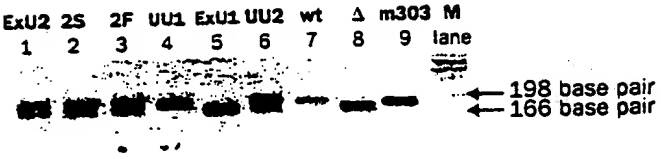
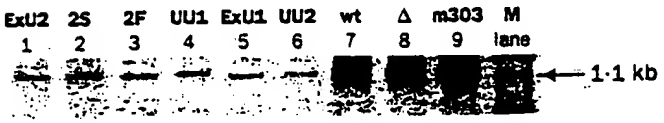

HeLa Ada
HeLa LAI
HeLa Δ envelope
b-galactosidase activity (OD)
2·5
2·0
1·5
1·0
0·5
0
CCR5wt
CCR5Δ32
CCR5m303

FIGURE 2


312
N T M C Q L L
aat aca atg tgt caa ctc ttg... wt
HincII
N T M Z Q L L
aat aca atg tga caa ctc ttg... m303
Absence of HincII


ExU2 2S 2F UU1 ExU1 UU2 wt Δ m303 M
1 2 3 4 5 6 7 8 9 lane
198 base pair
166 base pair


ExU2 2S 2F UU1 ExU1 UU2 wt Δ m303 M
1 2 3 4 5 6 7 8 9 lane
1·1 kb

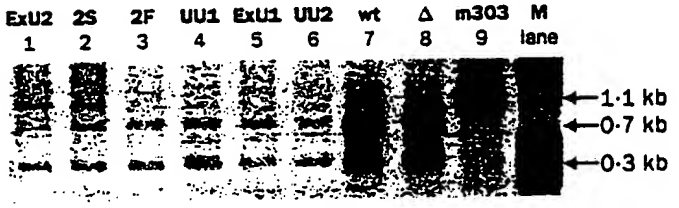

ExU2 2S 2F UU1 ExU1 UU2 wt Δ m303 M
1 2 3 4 5 6 7 8 9 lane
1·1 kb
0·7 kb
0·3 kb

FIGURE 3

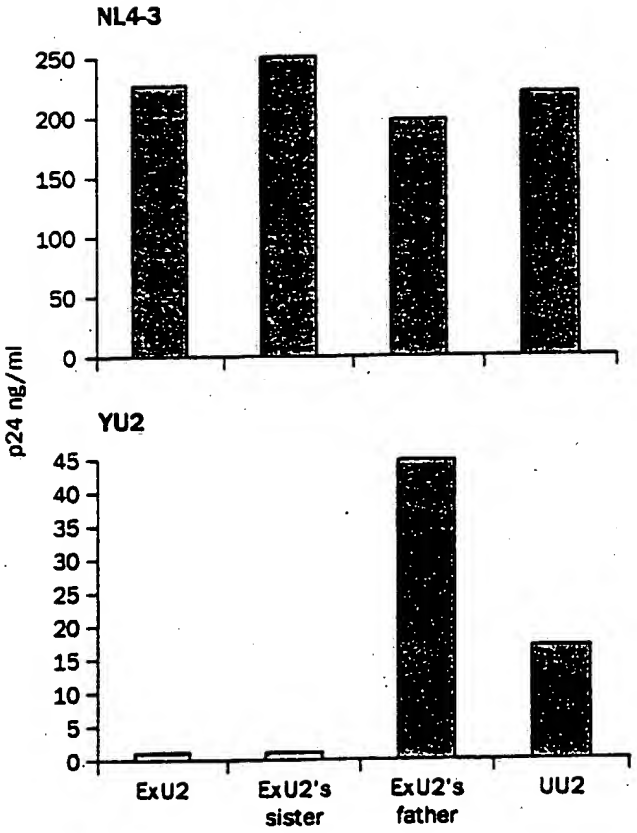
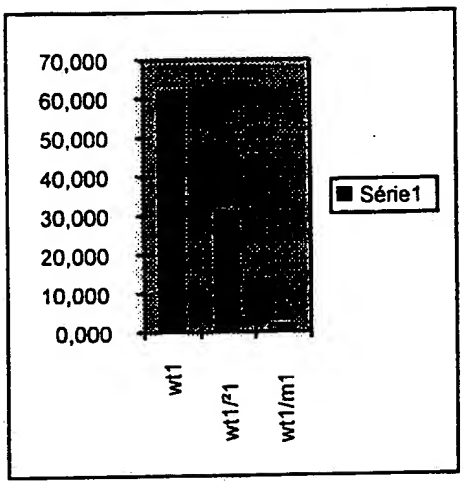
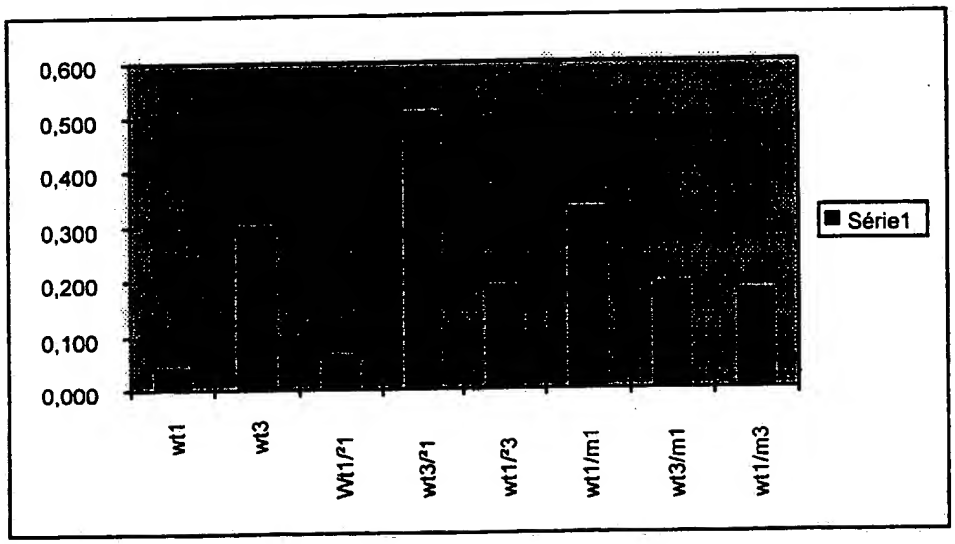

NL4-3
250
200
150
100
50
0
p24 ng/ml
YU2
45
40
35
30
25
20
15
10
5
0
ExU2
ExU2's sister
ExU2's father
UU2

FIGURE 4


70,000
60,000
50,000
40,000
30,000
20,000
10,000
0,000
Série1
wt1
wt1/r1
wt1/m1

A


0,600
0,500
0,400
0,300
0,200
0,100
0,000
Série1
wt1
wt3
Wt1/r1
wt3/r1
wt1/r3
wt1/m1
wt3/m1
wt1/m3

B

FIGURE 5

INTERNATIONAL SEARCH REPORT

Int. ional Application No
PCT/EP 98/03437

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C12N15/11 C12N15/62 C12N5/10 C07K14/715
C07K16/28 C12Q1/68 C12Q1/70 G01N33/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K C12Q G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 96 39437 A (HUMAN GENOME SCIENCES INC ;LI YI (US); RUBEN STEVEN M (US)) 12 December 1996 see the whole document ---	6,7,9, 10,12,21
X	WU L ET AL: "CCR5 LEVELS AND EXPRESSION PATTERN CORRELATE WITH INFECTABILITY BY MACROPHAGE-TROPIC HIV-1, IN VITRO" JOURNAL OF EXPERIMENTAL MEDICINE, vol. 185, no. 9, 5 May 1997, pages 1681-1691, XP002060735 see page 1683, right-hand column, line 25 - line 40 see page 1683, right-hand column, line 41 - left-hand column, line 27 see page 1685, left-hand column, paragraph 2 - page 1686, right-hand column, paragraph 1 --- -/--	6,7,9, 10,12, 21,26,27

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
3 November 1998	16/11/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Hornig, H

4

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 97 22698 A (ICOS CORP) 26 June 1997 see the whole document	6,7,9, 10,12,21
P,X	WO 97 32019 A (EUROSCREEN S A ;SAMSON MICHEL (FR); PARMENTIER MARC (BE); VASSART) 4 September 1997 see the whole document	6,7,9, 10,12, 21,26,27
P,X	WO 97 44055 A (UNIV NEW YORK) 27 November 1997 see the whole document	6,7,9, 10,12,21
P,X	WO 97 47319 A (PROGENICS PHARM INC ;AARON DIAMOND AIDS RESEARCH CE (US)) 18 December 1997 see the whole document	27
P,X	WO 97 45543 A (COMBADIERE CHRISTOPHE ;FENG YU (US); US HEALTH (US); ALKHATIB GHAL) 4 December 1997 see the whole document	27
P,X	WO 98 05798 A (AARON DIAMOND AIDS RESEARCH CE) 12 February 1998 see the whole document	6,7,9, 10,12,21
P,X	WO 98 18826 A (LEUKOSITE INC) 7 May 1998 see the whole document	26,27
P,X	QUILLENT C ET AL: "HIV-1-resistance phenotype conferred by combination of two separate inherited mutations of CCR5 gene 'see comments!." LANCET, (1998 JAN 3) 351 (9095) 14-8. JOURNAL CODE: LOS. ISSN: 0140-6736., XP002081596 see the whole document	1-7,9, 10,12, 19-24
P,X	GARRED P ET AL: "Chemokine-receptor polymorphisms: clarity or confusion for HIV-1 prognosis" THE LANCET, XP002082988 LANCET,LONDON,UK see the whole document	1-5

4

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int. ional Application No
PCT/EP 98/03437

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9639437 A	12-12-1996	AU 2663295 A	24-12-1996
		EP 0815137 A	07-01-1998
WO 9722698 A	26-06-1997	AU 1689297 A	14-07-1997
		BR 9607300 A	25-11-1997
		CA 2213331 A	26-06-1997
		CN 1183805 A	03-06-1998
		CZ 9702610 A	17-06-1998
		EP 0811063 A	10-12-1997
		HU 9801127 A	28-08-1998
		NO 973800 A	20-10-1997
		PL 321937 A	05-01-1998
WO 9732019 A	04-09-1997	NONE.	
WO 9744055 A	27-11-1997	AU 3143497 A	09-12-1997
WO 9747319 A	18-12-1997	AU 3390297 A	07-01-1998
		AU 3402697 A	07-01-1998
		WO 9747318 A	18-12-1997
WO 9745543 A	04-12-1997	AU 3375697 A	05-01-1998
WO 9805798 A	12-02-1998	AU 4055897 A	25-02-1998
WO 9818826 A	07-05-1998	AU 5155398 A	22-05-1998